Bayesian Estimation of Probabilistic Atlas for Anatomically-Informed Functional MRI Group Analyses

Hao Xu, Bertrand Thirion, Stéphanie Allassonnière

Bayesian Estimation of Probabilistic Atlas for Anatomically-Informed Functional MRI Group Analyses

Hao Xu, Bertrand Thirion, and Stéphanie Allassonnière

CMAP Ecole Polytechnique, Route de Saclay, 91128 Palaiseau, France
Parietal Team, INRIA Saclay-Île-de-France
CEA, DSV,I^2BM, Neurospin bât 145, 91191 Gif-Sur-Yvette, France

Abstract. Traditional analyses of Functional Magnetic Resonance Imaging (fMRI) use little anatomical information. The registration of the images to a template is based on the individual anatomy and ignores functional information; subsequently detected activations are not confined to gray matter (GM). In this paper, we propose a statistical model to estimate a probabilistic atlas from functional and T1 MRIs that summarizes both anatomical and functional information and the geometric variability of the population. Registration and Segmentation are performed jointly along the atlas estimation and the functional activity is constrained to the GM, increasing the accuracy of the atlas.

Keywords: Probabilistic atlas, geometric variability, joint registration segmentation, atlas-based segmentation, multi-modal, T1 MRI and fMRI

1 Introduction

Brain atlas is a useful tool in medical image analysis for both segmentation and registration. Probabilistic atlases yield a useful summary of a given dataset [6, 7], as they take into account the uncertainty on the underlying tissue type, which is related to partial volume effect (PVE) or to perfectible registration. In [3], a probabilistic framework was proposed for joint nonlinear registration, intensity normalization and segmentation of a single image, from which it infers tissue probability maps. In [10], a probabilistic model was proposed to segment a heterogeneous data set of brain MRIs simultaneously while constructing the probabilistic atlases. In spite of its convincing results, this model is not consistent as the deformations are considered as parameters (whereas segmentation is an unobserved random variable). In [13], the model proposed in [3] was generalized in order to provide estimates of individual segmentation as well as the probabilistic atlas from a set of anatomical images. This approach handles both the segmentation and registration as hidden variables, leading to a coherent convergent statistical estimator. However, this model is limited to scalar images. Here, we generalize it to create a probabilistic atlas that provides the probabilistic templates of each tissue as well as the degree of activation on GM voxels and the geometric variability.

Functional Magnetic Resonance Imaging of the brain is used to localize functional areas in the cortex and deep nuclei by measuring MRI signal changes associated with neural activity. It is a tool of choice for cognitive studies that aim at identifying specific regions of the brain that are activated in perceptual, cognitive or motor tasks. The most popular type of analysis is Statistical Parametric Mapping (SPM) [5], an approach that estimates the probability that some activation can be due to chance alone and provides p-value maps. Group analysis is then used to detect regions that show a positive mean activation across subjects [4, 12]. Accurate realignment of individual scans is most often obtained by normalizing individual anatomical images to a T1 MRI template. These processing steps are done without considering the complementarity of the anatomical and functional information available in each subject. Therefore, detected activations are not confined to gray matter. Few fMRI segmentation methods have been proposed to take into account multi-modal data, such as T1 and functional MRI. An implementation of cortical-based analysis of fMRI data was proposed in [2]. The fMRI data is mapped to the cortical surface, then activations are detected on the surface. It has been shown to achieve anatomically accurate activation detection. In [8], Markov Random Fields (MRF) were used as a spatial regularization in fMRI detection and anatomical information was incorporated into the MRF-based detection framework. In [11], both anatomical and functional data are used to improve the group-wise registrations. Anatomical information appears helpful in fMRI detection; however, the approaches so far do not incorporate a group model into the analysis. In this paper, we process multi-modal data jointly to ensure that the detected active areas are conditioned to gray matter while registration is informed by functional information. More specifically, group analysis first performs the realignment of individual images to a T1 MRI template and then segments active regions by thresholding. However, performing registration and segmentation jointly is generally more effective than performing them sequentially [13, 14]. In this paper, we take advantage of such coupling.

To deal with all the issues described above, we propose an atlas estimation procedure that can improve the template image estimation and the detection of the active areas. We generalize the model proposed in [13]. The input is now multivariate, as it encodes multi-modal patient observations (gray level T1 and functional MRIs). The estimated active areas are conditioned to GM segmentation. We perform the estimation by coupling the segmentation and registration steps. We estimate a probabilistic atlas that accounts for the variability of active areas in the population. We also learn the geometry as the metric on the space of deformations which drives the coupled segmentation. We use a stochastic algorithm with known guarantees on the convergence in the estimation procedure. The output of the algorithm is the probabilistic atlas, the individual active areas and the means and variances of each tissue type in each modality.

The rest of this paper is organized as follows. In Section 2, we present the model, the estimation procedure, the algorithm. Section 3 yields experimental results on simulated and real data.

2 Methods

Statistical Model. Let us consider n pairs of T1- and f- MRIs $(y_{1,i}, y_{2,i})_{1 \leqslant i \leqslant n}$ from n patients. Each image is observed on a grid of voxels Λ embedded in a continuous domain $D \subset \mathbb{R}^3$. We denote $x_j \in D$ the location of voxel j. We consider that each T1 MRI is composed of voxels belonging to one of the four classes, corresponding to four tissue types: gray matter, white matter (WM), CSF and background (BG). Each fMRI is composed of voxels belonging to one class among $3 + K$, corresponding to WM, CSF and BG, where no activation is expected to occur, and K different levels of activation in gray matter. We assume that the signal in the $3 + K$ classes is normally distributed with class dependent means $(\mu_{1,f(k)}, \mu_{2,k})_{k \in \{WM,CSF,BG,GM_1,\dots GM_K\}}$ and variances $(\sigma^2_{1,f(k)}, \sigma^2_{2,k})_{k \in \{WM,CSF,BG,GM_1,\dots GM_K\}}$, where $f(k) = k$ if $k \in \{WM, CSF, BG\}$, GM otherwise. The whole set of parameters is denoted by Θ.

As mentioned previously, we are working with gray level images which have not been pre-segmented. The unknown class of each voxel is supposed to be the discretization on Λ of a random deformation of probability maps $(P_k)_{1 \leqslant k \leqslant K+3}$. These probability maps correspond to the probability of each voxel to belong to each class in the template domain. They form the *probabilistic template* of the population. The random deformations from this template to each subject are also unobserved as the images are not pre-registered. We define them through a random field $z : \mathbb{R}^3 \to \mathbb{R}^3$ such that for $j \in \Lambda$ the prior probability of a voxel j from subject i to be in the k^{th} class is given by:

$$\mathbb{P}(c_i^j = k) = P_k(x_j - z(x_j))\,. \tag{1}$$

We define the deformation field as a finite linear combination of a given kernel K_g centered at some fixed equi-distributed control points in the domain D, $(x_g)_{1 \leqslant g \leqslant k_g}$, with parameter $\beta \in (\mathbb{R}^3)^{k_g}$

$$\forall x \in D, z_\beta(x) = \sum_{g=1}^{k_g} K_g(x, x_g)\beta(g)\,, \tag{2}$$

where K_g is chosen as a radial Gaussian Kernel. Note that we expect the tissue-specific information to be found in all the brain volume, hence the whole volume has to be covered with control points. As for the deformation model, the probability template maps $P_k : \mathbb{R}^3 \to]0, 1[, \forall k \in [\![1, K + 3]\!]$ are parametrized by the coefficients $\alpha_k \in]0, 1[^{k_p}$ which satisfy $\forall l \in [\![1, k_p]\!]$, $\sum_{k=1}^{K+3} \alpha_k^l = 1$. Let $(p_l)_{1 \leqslant l \leqslant k_p}$ be some control points :

$$\forall x \in D, P_k(x) = \sum_{l=1}^{k_p} K_p(x, p_l)\alpha_k^l, \tag{3}$$

where $K_p(x, p_l) = 1$ if p_l is the nearest neighbor of x among $(p_j)_j, 0$ otherwise.

The previous hypothesis provides a generative statistical model for a sample of pairs of gray level images. The random variables are the deformation vector β, the class of each voxel c and the observed gray levels of the images. We assume that the deformation vector follows a normal distribution with mean zero and non-diagonal covariance matrix Γ_g. The hierarchical model is given by:

$$\begin{cases} \beta_i \sim \mathcal{N}(0, \Gamma_g)|\Gamma_g; \quad c_i^j \sim \sum_{k=1}^{K+3} \delta_k P_k(x_j - z_{\beta_i}(x_j))|\beta_i, \\ \begin{pmatrix} y_{1,i}^j \\ y_{2,i}^j \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_{1,f(k)} \\ \mu_{2,k} \end{pmatrix}, \begin{pmatrix} \sigma_{1,f(k)}^2 & 0 \\ 0 & \sigma_{2,k}^2 \end{pmatrix} \right) \Big| c_i^j = k, \begin{pmatrix} \mu_{1,f(k)} \\ \mu_{2,k} \end{pmatrix}, \begin{pmatrix} \sigma_{1,f(k)}^2 \\ \sigma_{2,k}^2 \end{pmatrix} \end{cases} \tag{4}$$

where $\mathcal{N}(\cdot; \mu, \sigma^2)$ is the normal density with mean μ and variance σ^2 and δ_k is a Dirac function. The covariance matrix Γ_g is not assumed to have any particular pattern of zeros. This makes it possible to model local and global correlations between control point moves, in particular, very correlated displacements can be captured such as translation of a large area of the images.

The parameters to estimate are the covariance matrix Γ_g of the deformation distribution (Eq. (2)), $(\alpha_k)_{1 \leqslant k \leqslant K+3}$ the coefficients that define the template maps (Eq. (3)), $(\mu_{1,f(k)}, \mu_{2,k})_{1 \leqslant k \leqslant K+3}$ and $(\sigma^2_{1,f(k)}, \sigma^2_{2,k})_{1 \leqslant k \leqslant K+3}$ the class dependent means and variances. As medical images are high-dimensional but usually come in small samples, we work in a Bayesian framework. We use the standard conjugate priors for the covariance matrix, the class dependent means and variances with fixed hyper-parameters. All priors are assumed independent.

Estimation Algorithm. A maximum a posteriori (MAP) approach yields estimates of the model parameters: $\hat{\theta}_n = \arg\max_{\theta \in \Theta} q_B(\theta|(y_{1,1}, y_{2,1}), \cdots, (y_{1,n}, y_{2,n}))$, where q_B denotes the posterior distribution of the parameters given the n observations $(y_{1,1}, y_{2,1}), ..., (y_{1,n}, y_{2,n})$. As we are in an incomplete-data setting, we choose the Stochastic Approximation Expectation-Maximization (SAEM) algorithm coupled with a Markov Chain Monte Carlo method to take advantage of its theoretical and numerical properties [1, 13]. The SAEM algorithm is an iterative procedure that consists of three steps. First, we simulate the missing data using a Metropolis-Hastings algorithm within Gibbs sampler. Then a stochastic approximation is done on the sufficient statistics using the simulated value of the missing data. Last, we maximize the expected log-likelihood with respect to the model parameters.

3 Experiments and Results

We test our algorithm on both simulated data and real data. As the SAEM algorithm is an iterative procedure, we run 250 iterations which was checked to reach convergence. We initialize $\beta_0 = 0$ and a random classification c_0.

Simulated data. We use a pair of $64 \times 64 \times 8$ images as the reference images. We consider here $K = 3$, i.e. three different levels of activation in GM and 6

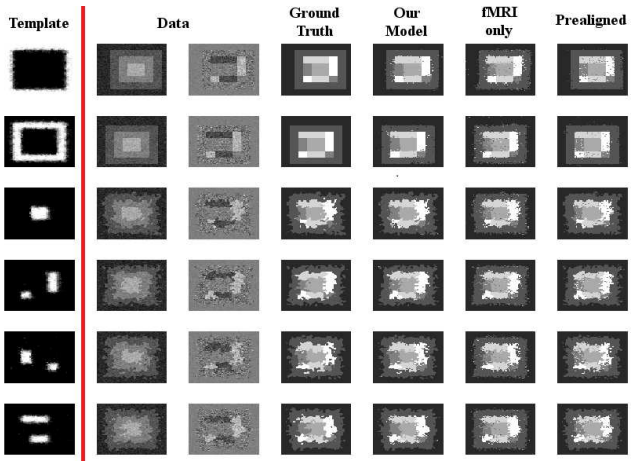

Fig. 1. Experiments on simulated data. The first column displays the first slice of the probabilistic template, each row corresponding to a class and white/black colors to high/low probability. The second and third columns show one slice of six pairs of data images. The forth to seventh columns correspond to the ground truth and the estimated segmentation for different models.

classes in total. We define the means and the standard deviations as follows (taking values that are observed in real fMRI for the standard deviations):

$$\begin{pmatrix}\mu_{1,1:6}\\ \mu_{2,1:6}\end{pmatrix} = \begin{pmatrix}1 & 2 & 4 & 3 & 3 & 3\\ 0 & 0 & 0 & 2.5 & 0 & -2.5\end{pmatrix}, \quad \begin{pmatrix}\sigma_{1,1:6}\\ \sigma_{2,1:6}\end{pmatrix} = \begin{pmatrix}0.25 & 0.25 & 0.25 & 0.25 & 0.25 & 0.25\\ 0.24 & 1.22 & 0.91 & 0.78 & 0.71 & 0.83\end{pmatrix}$$

The training data is composed of 20 pairs of images with random deformations of our template following Eq. (4) with previous parameters. We take 64 fixed control points for the deformation model given in Eq. (2), i.e. one control point in each $4 \times 4 \times 4$ cube. We take all the points in the image as landmarks for the template model given in Eq. (3).

The most important output of our estimation procedure is the probabilistic template. The estimated probabilistic maps are shown in the first column in Fig.1, each row corresponding to one class. The white/dark colors represent high/low probability of the tissues. Our probabilistic maps are sharp, as most voxels in each class have a probability larger than 0.9. Only voxels at the boundary between two classes are fuzzy which takes into account both the accuracy of deformation and the level of noise.

As mentioned previously, our model uses both the T1- and f- MRIs because we want the active areas to be conditioned to GM. We compare our model with the segmentation model in [13] using fMRI only. The result is shown in Fig.1. The second and third columns correspond to one slice of six pairs of data images. The ground truth and the final estimated segmentation of different methods (our model, the model using fMRI only) are shown from the forth to sixth columns. From the forth and fifth columns, we can see that the segmentation obtained

	BG	CSF	WM	GM_1	GM_2	GM_3
Our model	**98.5%**	**92.3%**	**88.4%**	**88.0%**	**73.1%**	**90.4%**
fMRI only	89.0%	75.0%	69.7%	78.3%	48.2%	81.4%
Pre-Aligned	96.6%	91.1%	87.0%	86.6%	67.8%	86.6%

Table 1. Experiments on synthetic data. Jaccard Index for the different methods averaged across all data.

with our atlas estimation is accurate. From the fifth and sixth columns, we see the improvement using the information provided from the T1 MRI. We calculate the Jaccard index for each class as a quantitative validation (Table 1) for each method. Our model yields an accurate segmentation, as only few voxels are misclassified. For the model using fMRI only, we are only interested in GM_1 and GM_3 which correspond to the active areas. As the other classes are non-active, the means of these classes are close to zero, therefore they are difficult to segment without the MRI tissue type information, leading to lower values for these classes. Moreover, the Jaccard indexes for the active area are lower than those obtained with our procedure, which shows that the coupling of both images information increases the accuracy of detection.

In our model, the registration and segmentation are done jointly, which avoids any pre-registration. In the preprocessing, each fMRI is pre-aligned to its corresponding MR image. However, the inter-subject non-rigid registration is not done, as it would require a template and would not take into account the fMRI observation to drive this preprocessing step. We compare our model with the pre-aligned model which does the registration and the segmentation sequentially. Fisrt, we use the segmentation model in [13] using the T1 MRIs, we get the deformation vector and individual tissue segmentation as our output. Then we apply the same deformation to the fMRI and detect the activation only in GM. The estimated segmentation of the pre-aligned model is shown in the seventh columns in Fig.1. Comparing the fifth and seventh columns, the segmentations look similar which makes it difficult to say which method gives the better result. However our model gives less isolated points. Moreover, looking at the Jaccard indexes (Table 1), we see that our model outperforms the pre-aligned model. This shows the improvement of doing registration and segmentation jointly.

In-vivo data. The proposed method was also tested on a real MRI and fMRI dataset described in [9]. Both anatomical and functional data were subject to standard preprocessing using SPM8, including spatial normalization and General Linear Model analysis. Images are sampled at 3mm resolution, yielding volumes of shape $46 \times 53 \times 63$. We select a contrast from the fMRI that yields differential effect of a computation task versus a simple instruction reading/listening. We have $K = 3$ levels of activation in the GM, hence 6 classes overall.

We take 792 fixed control points for the deformation model given in Eq. (2), corresponding to one control points in each $6 \times 6 \times 6$ cube and $23 \times 27 \times 32$ points in the image as the landmarks for the template model given in Eq. (3), corresponding to one landmark in each $2 \times 2 \times 2$ cube.

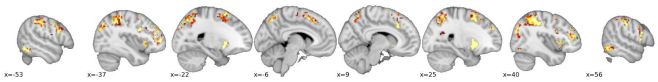

Fig. 2. The estimated template on real data. The yellow/red colors correspond to high/low probability of the activation for the computation task.

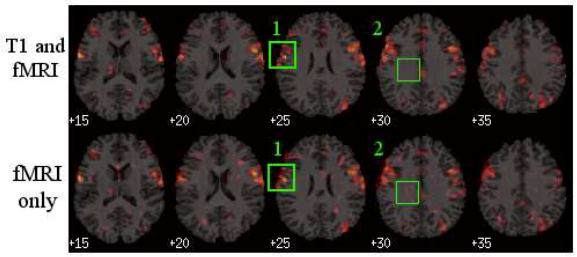

Fig. 3. Experiments on real data showing the detected active areas $p > .95$ for the computation task. The first row for our method using both T1- and f- MRI and the second row for the standard method using fMRI only. Each column corresponds to one slice of the same patient.

The estimated probabilistic maps, thresholded at the $p > .95$ level, are shown in Fig.2. The yellow/red colors correspond to high/low probability of the computation task activation. Our probabilistic maps are sharp. The detected areas are well conditioned to GM and fits the known active areas for the computation task. For example, in the slice $x = 25mm$, we find well the Putamen.

We compared our model with the standard method that thresholds the group-level mean activation. We represent the active areas in the computation task overlayed on T1 images. The results of one patient are shown in Fig.3. The first row for our method uses both T1- and f- MRI and the second row for the method uses fMRI only. Each column corresponds to one slice of the same patient. In zone 1, we see that the areas detected as active by our method are limited to the GM. However, a part of the detected active areas by the non-anatomically aware method are outside of the brain. In zone 2, the standard method detects some active areas in WM, while our method does not. These show that we reach our goal, i.e. the detected active areas are well conditioned to GM. The detected active areas by our method are similar to those by the standard method in GM, this shows that our segmentation is accurate.

4 Conclusion

In this study, we proposed a statistical model to detect the active areas in the brain using both T1 and functional MRI. We used a stochastic algorithm to perform registration, segmentation and to create a probabilistic atlas simultaneously. Our model has several advantages. First, the probabilistic atlas contains

both the templates and the geometric variability of the population. Second, we do not need any pre-registration to perform the segmentation which is automatically obtained as an output. Third, the detected active areas are confined to GM with the information provided from the MRI data. Our experiments show that we get better results with our algorithm than the standard method. The detected active areas are well conditioned to GM and the atlas is sharp.

Acknowledgment

We thank Digiteo for funding MMoVNI.

References

1. Allassonnière, S., Kuhn, E., Trouvé, A.: Construction of bayesian deformable models via stochastic approximation algorithm: A convergence study. Bernoulli Journal 16(3), 641–678 (2010)
2. Andrade, A., Kherif, F., Mangin, J.F., Worsley, K.J., Paradis, A.L., Simon, O., Dehaene, S., Bihan, D.L., Poline, J.B.: Detection of fMRI activation using cortical surface mapping. Human Brain Mapping 12, 79–93 (2001)
3. Ashburner, J., Friston, K.J.: Unified segmentation. NeuroImage 26(3), 839–851 (2005)
4. Friston, K.J., Holmes, A.P., Price, C.J., Büchel, C., Worsley, K.J.: Multisubject fMRI studies and conjunction analyses. NeuroImage 10(4), 385–396 (1999)
5. Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.B., Firith, C.D., Frackowiak, R.S.J.: Statistical prametric maps in functional imaging: a general linear approach. Human Brain Mapping 2(4), 189–210 (1995)
6. Gouttard, S., Styner, M., Joshi, S., Smith, R.G., Cody, H., Gerig, G.: Subcortical structure segmentation using probabilistic atlas priors. SPIE (2007)
7. Leemput, K.V., Maes, F., Vandermeulen, D., Colchester, A., Suetens, P.: Automated segmentation of multiple sclerosis lesions by model outlier detection. IEEE TMI 20(8), 677–688 (2001)
8. Ou, W., Wells III, W.M., Golland, P.: Combining spatial priors and anatomical information for fMRI detection. Medical Image Analysis 14(3), 318–331 (2010)
9. Pinel, P., Thirion, B., Meriaux, S., Jobert, A., Serres, J., Bihan, D.L., Poline, J.B., Dehaene, S.: Fast reproducible identification and large-scale databasing of individual functional cognitive networks. BMC Neuroscience 8 (2007)
10. Ribbens, A., Hermans, J., Maes, F., Vandermeulen, D., Suetens, P.: Sparc: Unified framework for automatic segmentation, probabilistic atlas construction, registration and clustering of brain mr images. IEEE ISBI pp. 856–859 (2010)
11. Sabuncu, M.R., Singer, B.D., Conroy, B., Bryan, R.E., Ramadge, P.J., Haxby, J.V.: Function-based intersubject alignment of human cortical anatomy. Cerebral Cortex 20, 130–140 (2010)
12. Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., Poline, J.B.: Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. NeuroImage 35, 105–120 (2007)
13. Xu, H., Thirion, B., Allassonnière, S.: Probabilistic atlas and geometric variability estimation to drive tissue segmentation. Statistics in Medicine (submitted)
14. Yeo, B., Sabuncu, M.R., Vercauteren, T., Ayache, N., Fischl, B., Golland, P.: Spherical demons: Fast diffeomorphic landmark-free surface registration. IEEE TMI 29(3), 650–668 (2010)